

534,175

Rec'd PC

05 MAY 2005

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2004 年 5 月 27 日 (27.05.2004)

PCT

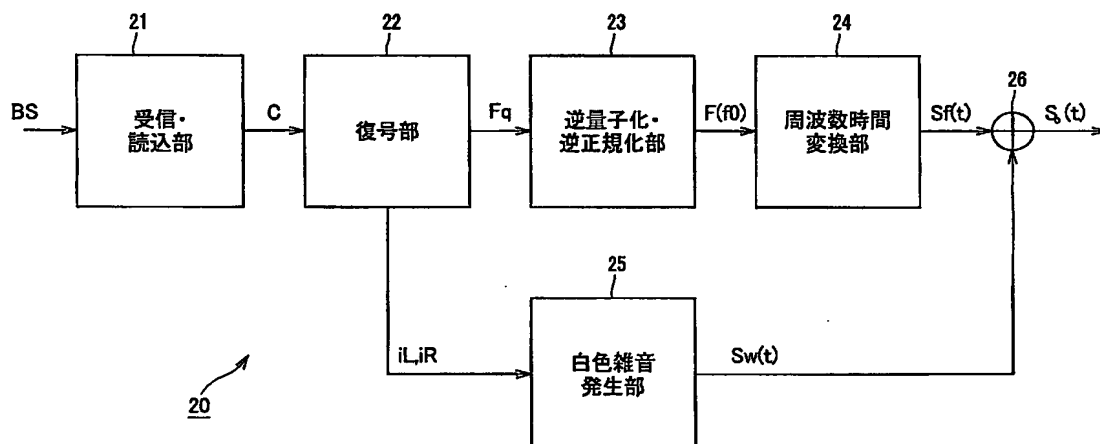
(10) 国際公開番号
WO 2004/044891 A1

- (51) 国際特許分類: G10L 19/02 (72) 発明者; および
(21) 国際出願番号: PCT/JP2003/013084 (75) 発明者/出願人 (米国についてのみ): 鈴木 志朗 (SUZUKI,Shiro) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP). 辻 実 (TSUJI,Minoru) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP). 東山 恵祐 (TOYAMA,Keisuke) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).
(22) 国際出願日: 2003 年 10 月 10 日 (10.10.2003)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(30) 優先権データ:
特願 2002-330024
2002 年 11 月 13 日 (13.11.2002) JP (74) 代理人: 小池 晃, 外(KOIKE,Akira et al.); 〒100-0011 東京都千代田区 内幸町一丁目 1 番 7 号 大和生命ビル 1 1 階 Tokyo (JP).
(71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川 6 丁目 7 番 3 5 号 Tokyo (JP). (81) 指定国 (国内): CN, KR, US.

[続葉有]

(54) Title: MUSIC INFORMATION ENCODING DEVICE AND METHOD, AND MUSIC INFORMATION DECODING DEVICE AND METHOD

(54) 発明の名称: 音楽情報符号化装置及び方法、並びに音楽情報復号装置及び方法



21...RECEPTION/READ-IN SECTION
22...DECODING SECTION
23...DE-QUANTIZATION/DE-NORMALIZATION SECTION
24...FREQUENCY TIME CONVERSION SECTION
25...WHITE NOISE GENERATION SECTION

(57) Abstract: In a music information encoding device, when encoding a music signal containing a white noise, a code string includes an index iL of the energy level of the white noise component and an index iR specifying a random number table start index. In a music information decoding device (20), a white noise generation section (25) generates a white noise signal $S_w(t)$ on the temporal axis having the level equivalent to the white noise by using the index iL and iR contained in the code string C and an adder (26) adds the decoded sound signal $S_f(t)$ on the temporal axis and outputs an output music signal $S_0(t)$.

(57) 要約: 音楽情報符号化装置では、白色雑音成分を含む音楽信号を符号化する際に、白色雑音成分のエネルギーレベルのインデックス iL と乱数テーブルの開始インデ

[続葉有]

WO 2004/044891 A1



(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

ックスを指定するインデックス i_R を符号列に含める。音楽情報復号装置 (20) では、白色雑音発生部 (25) において、符号列 C に含まれたインデックス i_L 、 i_R を用いてその白色雑音と同等のレベルをもつ時間軸上の白色雑音信号 $S_w(t)$ を発生させ、加算器 (26) において、復号した時間軸上の音楽信号 $S_f(t)$ と加算し、出力音楽信号 $S_o(t)$ として出力する。

明細書

音楽情報符号化装置及び方法、並びに音楽情報復号装置及び方法

技術分野

本発明は、白色雑音成分を含む音楽情報を符号化する音楽情報符号化装置及びその方法、この音楽情報符号化装置及び方法によって生成された符号列の記録された記録媒体、この音楽情報符号化装置及び方法によって生成された符号列を復号する音楽情報復号装置及びその方法、並びにこの音楽情報符号化処理又は音楽情報復号処理をコンピュータに実行させるプログラムに関する。

本出願は、日本国において2002年11月13日に出願された日本特許出願番号2002-330024を基礎として優先権を主張するものであり、この出願は参照することにより、本出願に援用される。

背景技術

従来より、入力音楽信号を符号化する際には、時間軸上の音楽信号を一定の時間区間（フレーム）毎にブロック化し、フレーム毎に改良離散コサイン変換（Modified Discrete Cosine Transformation; MDCT）等を行うことで、時間軸上の時系列信号を周波数軸上のスペクトル信号に変換（スペクトル変換）して符号化することが行われている。

また、スペクトル信号を符号化する際には、フレーム毎の時系列信号をスペクトル変換したスペクトル信号毎に所定のビット配分、或いは適応的なビット割当（ビットアロケーション）が行われる。すなわち、例えば、MDCT処理されて得られた係数データをビットアロケーションによって符号化する際には、ブロック毎の時間軸信号をMDCT処理して得られるMDCT係数データに対して、適応的にビット数が割り当てられて符号化が行われる。

なお、このビットアロケーションについては、例えば、文献「音声信号の適応

変換符号化」(“Adaptive Transform Coding of Speech Signals”, R. Zelinski and P. Noll, IEEE Transactions of Acoustics, Speech and Signal Processing, vol. ASSP-25, No. 4, August 1977) や、文献「臨界帯域符号化 —聴覚システムの知覚の要求に関するディジタル符号化」(ICASSP 1980, “The critical band coder digital encoding of the perceptual requirements of the auditory system”, M. A. Kransner MIT) 等にその詳細が記載されている。

ところで、符号化装置への入力音楽信号には、楽器、声等の様々な成分が存在している。例えば、声やピアノの音のみをマイクロホンにて録音した場合においても、純粹にそれらの音のみが記録されている訳ではなく、背景雑音や録音機器の動作音、或いは録音機器自体の電氣的雑音が多少なりとも記録されるのが普通である。

符号化装置からみれば、それらの雑音も声もピアノの音も1次元の波形情報でしかなく、雑音成分をも周波数変換して符号化しようとする。これは、波形再現性という観点からは正しいアプローチであるが、人間の聴覚特性を考慮した場合には効率的な符号化手法とはいえない。

そこで、聴覚心理モデルに基づくビットアロケーションによって、例えば絶対に聞こえないレベルである最低可聴レベル又は符号化装置にて任意に設定できる最低符号化閾値よりも小さい周波数成分に対してビット割当を行わないようにすることができる。

このようなビットアロケーションを行う従来の符号化装置の概略構成を図1に示す。図1に示すように、符号化装置100において、時間周波数変換部101は、入力音楽信号 $S_i(t)$ をスペクトル信号 $F(f)$ に変換し、このスペクトル信号をビット配分周波数帯域決定部102に供給する。ビット配分周波数帯域決定部102は、スペクトル信号 $F(f)$ を分析し、ビット割当を行う周波数成分、すなわち最低可聴レベル又は最低符号化閾値以上である周波数成分 $F(f_0)$ と、ビット割当を行わない周波数成分 $F(f_1)$ とに分割し、周波数成分 $F(f_0)$ のみを正規化・量子化部103に供給し、周波数成分 $F(f_1)$ を切り捨てる。

正規化・量子化部103は、周波数成分 $F(f_0)$ に対して正規化及び量子化を施し、生成された量子化値 F_q を符号化部104に供給する。符号化部104

は、この量子化値 F_q を符号化して符号列 C を生成し、記録・伝送部 105 は、この符号列 C を図示しない記録媒体に記録し、又はビットストリーム BS として伝送する。

この符号化装置 100 で生成される符号列 C の一例を図 2 に示す。図 2 に示すように、符号列 C は、ヘッダ H 、正規化情報 SF 、量子化精度情報 WL 及び周波数情報 SP からなる。

続いて、符号化装置 100 に対応する復号装置の概略構成を図 3 に示す。図 3 に示すように、復号装置 120 において、受信・読込部 121 は、符号化装置 100 から受信したビットストリーム BS 又は図示しない記録媒体から符号列 C を復元し、この符号列 C を復号部 122 に供給する。復号部 122 は、符号列 C を復号して量子化値 F_q を生成し、逆量子化・逆正規化部 123 は、この量子化値 F_q に逆量子化、逆正規化を施し、周波数成分 $F(f_0)$ を生成する。そして、周波数時間変換部 124 は、この周波数成分 $F(f_0)$ を出力音楽信号 $S(t)$ に変換して出力する。

ここで、符号化装置において、全てのフレームで最低可聴レベル A 未満の周波数成分に対してビット割当を行わないようにする場合の一例を図 4 に示す。図 4 に示すように、 $(n-1)$ 番フレームにおいては $0.60f$ 以下の周波数成分のみが符号化され、 n 番フレームにおいては $1.00f$ までの全ての周波数成分が符号化され、 $(n+1)$ 番フレームにおいては、 $0.55f$ 以下の周波数成分のみが符号化されることになる。この結果、フレームによって特定の周波数が符号列に含まれたり含まれなかったりするが、この符号列に含まれない周波数は人間の聴覚上、絶対的に聞こえないものであるため、全てのフレームにおいて全ての周波数成分を符号列に含めることと等価であり、後に再生した場合に聴覚心理的な違和感は生じない。

但し、このように最低可聴レベル以上の周波数成分を全て符号化する場合、本来重要でない周波数成分や聞こえなくともよい白色雑音まで符号化されるため、非効率的である。また、各フレームに同一のビット数を割り当てる固定ビットレートの符号化を行う場合には、ビットレートが低くなるに従って、満足な音質を達成するために必要なビット数を確保することができないフレームが出てくる虞

がある。

一方、符号化装置において、フレーム毎に設定された最低符号化閾値 a 未満の周波数成分に対してビット割当を行わないようにする場合の一例を図 5 に示す。図 5 に示すように、 $(n-1)$ 番フレームでは、符号化装置によって決定される最低符号化閾値が $a(n-1)$ というレベルに設定されている。この $a(n-1)$ という最低符号化閾値は、この値より小さい周波数であれば音質上それほど重要な成分でないため、 $(n-1)$ 番フレーム中においては記録しなくとも音質に与える影響は少ないと判定されるような値である。この結果、 $(n-1)$ 番フレームにおいては $0.60f$ 以下の周波数成分のみが符号化される。

このような符号化されない周波数成分が全てのフレームで一定であれば、低域通過フィルタを通してから全ての周波数成分を符号化するのとはほぼ等価であるため、聴覚上は帯域感が狭まるように感じる場合があるが、元の周波数分布と聴覚特性とを考慮すれば、狭帯域感は大きな問題にはならない。

しかしながら、続く n 番フレームでは全体のエネルギーが低いため、 $(n-1)$ 番フレームよりも符号化しない周波数成分が増えている。また、 $(n+1)$ 番フレームでは全体のエネルギーが高いため、符号化装置において全ての周波数成分が聴覚上重要であると判定され、全ての周波数成分が符号化されている。

このように、符号列に含める周波数成分がフレーム間で変動すると、後に再生する際に周波数成分のフレーム間の連続性がなくなり、明らかな聴覚上の雑音を感じることもある。その雑音は、FM 放送の背景雑音が電波状況の変動によって刻々と変化するようなものに似ており、音楽以外に一定の変調雑音が加算されているような感覚を受け、聴覚心理的な違和感が生じる。

そこで、本件出願人が先に提案した日本公開特許公報平 8-166799 号では、先行するフレームにおいてビット割当を行った帯域幅を記憶保持し、その帯域幅から大きく変動しないようにして現在のフレームにおいてビット割当を行う帯域幅を決定することにより、再生帯域の変動を抑制し、雑音の発生を防止する技術が開示されている。

しかしながら、この日本公開特許公報平 8-166799 号に記載の技術は、再生帯域の安定化に寄与するとはいえ、再生帯域の変動自体は許可しているため、

聴覚上の問題を完全に解決するものではない。

また、再生帯域を安定化するために、本来不必要と判定された帯域の周波数が記録されたり、本来必要と判定された帯域の周波数が記録されなかったりするため、符号化効率の観点から不利な場合も生じ得る。

この他に、数フレーム又は数十フレームに亘って全ての周波数を分析し、ビット割当を行う周波数を全てのフレーム間で揃えるということも考えられるが、実時間処理や民生用ハードウェアにおけるメモリ・プロセッサのコストを考慮すると実現は困難であり、また、符号化効率の向上も見込めない。

発明の開示

本発明は、このような従来の実情に鑑みて提案されたものであり、白色雑音成分を含む音楽情報を効率的に符号化すると共に、フレーム間での再生帯域の変動による雑音の発生を防止する音楽情報符号化装置及びその方法、この音楽情報符号化装置及び方法によって生成された符号列の記録された記録媒体、この音楽情報符号化装置及び方法によって生成された符号列を復号する音楽情報復号装置及びその方法、並びにこの音楽情報符号化処理又は音楽情報復号処理をコンピュータに実行させるプログラムを提供することを目的とする。

上述した目的を達成するために、本発明に係る音楽情報符号化装置及びその方法は、時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号化する際に、音楽信号中の白色雑音成分を分析し、分析した白色雑音成分のエネルギーレベルを表すインデックスを符号化する。

ここで、ブロック内の高域側のエネルギー分布に基づいて白色雑音成分を分析するようにしてもよく、ブロック全体のエネルギー分布に基づいて白色雑音成分を分析するようにしてもよい。

また、復号側で白色雑音成分を生成するために用いる乱数テーブルのインデックスをさらに符号化することもできる。

また、上述した目的を達成するために、本発明に係る記録媒体は、時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号

化すると共に、上記音楽信号中の白色雑音成分を分析し、該白色雑音成分のエネルギーレベルを表すインデックスを符号化して生成された符号列が記録されたものである。

また、上述した目的を達成するために、本発明に係る音楽情報復号装置及びその方法は、符号化された周波数信号を復号し、逆周波数変換して時間軸上の音楽信号を生成する際に、符号化された白色雑音成分のエネルギーレベルを表すインデックスに基づいて、時間軸上の白色雑音成分を生成し、逆周波数変換して得られる時間軸上の音楽信号と時間軸上の白色雑音成分とを加算する。

ここで、符号化された乱数テーブルのインデックスに基づいて白色雑音成分を生成するようにしてもよく、符号列中の所定の値に基づいて白色雑音成分を生成するようにしてもよい。

このような音楽情報符号化装置及びその方法、並びに音楽情報復号装置及びその方法では、白色雑音成分を含む音楽信号を符号化する際に、符号化側において白色雑音成分のエネルギーレベルのインデックスを符号列に含め、復号側においてその白色雑音と同等のレベルをもつ白色雑音を発生させ、復号した音楽信号と時間軸上で加算する。

また、本発明に係るプログラムは、上述した音楽情報符号化処理又は音楽情報復号処理をコンピュータに実行させるものである。

本発明のさらに他の目的、本発明によって得られる具体的な利点は、以下に説明される実施例の説明から一層明らかにされるであろう。

図面の簡単な説明

図 1 は、従来の符号化装置の概略構成を説明する図である。

図 2 は、同符号化装置で生成される符号列の一例を示す図である。

図 3 は、従来の復号装置の概略構成を説明する図である。

図 4 は、同符号化装置において、最低可聴レベル未満の周波数成分に対してビット割当を行わない場合の例である。

図 5 は、同符号化装置において、最低符号化閾値未満の周波数成分に対してビ

ット割当を行わない場合の例である。

図 6 は、符号化側における各フレームの最低符号化閾値及び白色雑音レベルの一例を示す図である。

図 7 は、復号側で生成される白色雑音の一例を示す図である。

図 8 は、本実施の形態における音楽情報符号化装置の概略構成を説明する図である。

図 9 は、インデックス i_L を生成するための白色雑音レベルテーブルの一例を示す図である。

図 10 は、インデックス i_R を生成するための乱数インデックステーブルの一例を示す図である。

図 11 は、同音楽情報符号化装置で生成される符号列の一例を示す図である。

図 12 は、本実施の形態における音楽情報復号装置の概略構成を説明する図である。

発明を実施するための最良の形態

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、白色雑音成分を含む音楽情報を効率的に符号化すると共に、再生帯域の時間的な変動による雑音の発生を防止する音楽情報符号化装置及びその方法、並びにこの音楽情報符号化装置及び方法によって生成された符号列を復号する音楽情報復号装置及びその方法に適用したものである。以下では、先ず、本実施の形態における音楽情報符号化方法及び音楽情報復号方法の原理について説明し、次いで本実施の形態における音楽情報符号化装置及び音楽情報復号装置の構成について説明する。

本実施の形態における音楽情報符号化方法では、時間軸上の入力音楽信号を一定の時間区間（フレーム）毎にブロック化し、フレーム毎に改良離散コサイン変換（Modified Discrete Cosine Transformation; MDCT）等を行うことで、時間軸上の時系列信号を周波数軸上のスペクトル信号に変換（スペクトル変換）して符号化する。この際、人間の聴覚特性を考慮して効率的に符号化するために、

聴覚心理モデルに基づくビットアロケーションによって、フレーム毎に設定可能な最低符号化閾値 a よりも小さい周波数成分に対してビット割当を行わないものとする。

例えば図 6 に示すように、 $(n-1)$ 番フレームでは、最低符号化閾値 a が $a(n-1)$ というレベルに設定される。この $a(n-1)$ という最低符号化閾値は、この値より小さい周波数であれば音質上それほど重要な成分でないため、 $(n-1)$ 番フレーム中においては記録しなくとも音質に与える影響は少ないと判定されるような値である。この結果、 $(n-1)$ 番フレームにおいては $0.60 f$ 以下の周波数成分に対してのみビット割当が行われる。

続く n 番フレームでは、最低符号化閾値 a が $a(n)$ というレベルに設定され、 $0.50 f$ 以下の周波数成分に対してのみビット割当が行われる。

また、 $(n+1)$ 番フレームでは、最低符号化閾値 a が $a(n+1)$ というレベルに設定され、 $1.0 f$ までの全ての周波数成分に対してビット割当が行われる。

ここで、最低符号化閾値 a 未満の周波数成分を切り捨てて符号列に含めない場合には、後に再生する際の再生帯域がフレーム間で変動し、フレーム間の連続性がなくなるため、聴覚心理的な違和感が生じてしまう。

そこで、本実施の形態では、最低符号化閾値 a 未満である高域側の周波数成分から白色雑音成分を分析し、

(a) 領域内のエネルギー分布が十分小さく、かつ平坦である。

(b) 領域内の周波数成分がノイズ性である。

という 2 つの条件を満たす領域の平均エネルギーレベルを量子化したインデックスを符号列に含める。

なお、ある領域内の周波数分布が平坦であり、周波数成分の最大値 f_{max} と平均値 f_{ave} との比 (f_{max}/f_{ave}) が 3.0 程度以下の場合に、その領域の周波数成分には周期性がなく、ノイズ性といえることが経験的に分かっている。

図 6 の例では、 $(n-1)$ 番フレーム、 n 番フレーム及び $(n+1)$ 番フレームについて、それぞれ高域の平坦な周波数のエネルギーレベルに一致するような白

色雑音レベル $b(n-1)$ 、 $b(n)$ 、 $b(n+1)$ を検出し、それらをインデックス化して符号列に含める。

一方、本実施の形態における音楽情報復号方法では、符号列に含まれた周波数成分をフレーム毎に時間軸上の信号に逆スペクトル変換して復号すると共に、インデックスが示すエネルギーレベルの白色雑音を発生させる。

この結果、図7に示すように、符号列に含まれた周波数成分の再生帯域はフレーム間で変動するものの、白色雑音によって擬似的に高域まで周波数を発生させることで、聴覚上の違和感を効果的に抑制することが可能となる。

なお、符号化側で符号列に含めないと判定された周波数成分のエネルギーレベルと、復号側で発生させた白色雑音のエネルギーレベルにはギャップがあるが、聴覚上の違和感の主たる原因は、ある周波数帯域のエネルギーが全くなってしまうことであるため、そのギャップが聴覚上悪影響を与えるようなことはない。

以上のような処理を行う本実施の形態における音楽情報符号化装置の概略構成を図8に示す。図8に示すように、音楽情報符号化装置10において、時間周波数変換部11は、入力音楽信号 $S_1(t)$ をスペクトル信号 $F(f)$ に変換し、このスペクトル信号 $F(f)$ をビット配分周波数帯域決定部12に供給する。

ビット配分周波数帯域決定部12は、スペクトル信号 $F(f)$ を分析し、ビット割当を行う周波数成分、すなわち最低符号化閾値 a 以上である周波数成分 $F(f_0)$ と、ビット割当を行わない周波数成分 $F(f_1)$ とに分割する。そして、ビット配分周波数帯域決定部12は、周波数成分 $F(f_0)$ を正規化・量子化部13に供給し、周波数成分 $F(f_1)$ を白色雑音レベル決定部14に供給する。

正規化・量子化部13は、周波数成分 $F(f_0)$ に対して正規化及び量子化を施し、生成された量子化値 F_q を符号化部15に供給する。

白色雑音レベル決定部14は、周波数成分 $F(f_1)$ から白色雑音成分を分析し、上述した2つの条件を満たす領域の平均エネルギーレベル、すなわち白色雑音レベルを量子化したインデックス i_L を生成する。このインデックス i_L を3ビットで表す場合、インデックス i_L を生成するための白色雑音レベルテーブルは、例えば図9に示すようになる。この例では、白色雑音レベルが約8 dBである場合、インデックス i_L は3となる。

また、白色雑音レベル決定部 14 は、復号側で白色雑音を発生させるために必要な乱数テーブルの開始インデックス i_{RT} を指定するためのインデックス i_R を生成する。このインデックス i_R を 3 ビットで表す場合、インデックス i_R を生成するための乱数インデックステーブルは、例えば図 10 に示すようになる。

符号化部 15 は、正規化・量子化部 13 から供給された量子化値 F_q と、白色雑音レベル決定部 14 から供給されたインデックス i_L 、 i_R とを符号化して符号列 C を生成し、記録・伝送部 16 は、この符号列 C を図示しない記録媒体に記録し、又はビットストリーム BS として伝送する。

この音楽情報符号化装置 10 で生成される符号列 C の一例を図 11 に示す。図 11 に示すように、符号列 C は、ヘッダ H 、正規化情報 SF 、量子化精度情報 WL 、及び周波数情報 SP の他に、白色雑音フラグ FL 及び白色雑音情報 WN からなる。また、白色雑音情報 WN は、インデックス i_L 及びインデックス i_R からなる。ここで、白色雑音フラグ FL が “1” の場合、白色雑音情報 WN が符号列 C に含まれる。一方、白色雑音フラグ FL が “0” の場合、白色雑音情報 WN は符号列 C に含まれず、余ったビットは周波数成分 $F(f_0)$ の符号化にまわされる。

なお、白色雑音フラグ FL を設けず、例えばフレーム内の全ての周波数成分が最低符号化閾値 a 以上である場合には、前フレームのインデックス i_L 、 i_R を符号列 C に含めるようにしても構わない。

続いて、音楽情報符号化装置 10 に対応する音楽情報復号装置の概略構成を図 12 に示す。図 12 に示すように、音楽情報復号装置 20 において、受信・読込部 21 は、音楽信号符号化装置 10 から受信したビットストリーム BS 又は図示しない記録媒体から符号列 C を復元し、この符号列 C を復号部 22 に供給する。

復号部 22 は、符号列 C を復号して量子化値 F_q とインデックス i_L 、 i_R とを生成し、量子化値 F_q を逆量子化・逆正規化部 23 に供給すると共に、インデックス i_L 、 i_R を白色雑音発生部 25 に供給する。

逆量子化・逆正規化部 23 は、量子化値 F_q に逆量子化、逆正規化を施して周波数成分 $F(f_0)$ を生成し、この周波数成分 $F(f_0)$ を周波数時間変換部 24 に供給する。

周波数時間変換部 24 は、この周波数成分 $F(f_0)$ を時間軸上の音楽信号 $S_r(t)$ に変換し、この音楽信号 $S_r(t)$ を加算器 26 に供給する。

白色雑音発生部 25 は、インデックス i_L , i_R から、以下の式 (1) に従って周波数成分 $F(f_1)$ に相当する時系列信号である白色雑音信号 $S_w(t)$ を発生し、この白色雑音信号 $S_w(t)$ を加算器 26 に供給する。

$$S_w(t) = LEV(i_L) * RND(i_{RT} + t) \quad \dots (1)$$

式 (1) において、 $LEV(i_L)$ は、インデックス i_L を引数とする白色雑音レベルテーブル $LEV()$ の値を示し、符号化側と共通の値である。また、 $RND(i_{RT} + t)$ は、乱数インデックステーブルにおいてインデックス i_R で指定される開始インデックス i_{RT} に周波数成分番号 t を加えた値を引数とする乱数テーブル $RND()$ の値を示す。この乱数テーブル $RND()$ の値は、例えば -1.0 以上 1.0 以下に正規化されている。

このように、符号列中のインデックス i_R により乱数テーブルの開始インデックス i_{RT} を生成することで、毎回異なる白色雑音が生成されることを防止することができる。

ここで、乱数テーブル $RND()$ では、 $i_{RT} + t$ の値が配列数 N_{rnd} を超える場合がある。このような場合には、例えば $i_{RT} + t$ から配列数 N_{rnd} を減算した値を乱数テーブル $RND()$ の引数とする。つまり $i_{RT} + t$ の値は 0 以上 N_{rnd} 以下としなければならない。

なお、本実施の形態では、符号列中のインデックス i_R により乱数テーブルの開始インデックス i_{RT} を生成するものとしたが、これに限定されるものではなく、例えば符号化側でインデックス i_R を生成せず、符号列中の所定の値、例えば 1 フレーム分の正規化情報 S_F 又は量子化精度情報 WL を全て加算した値に基づいて開始インデックス i_{RT} を生成するなどのようにしても構わない。この場合にも、毎回異なる白色雑音が生成されることを防止することができる。

また、毎回異なる白色雑音が生成されることを許容する場合には、復号側で乱数を発生させて開始インデックス i_{RT} を生成するようにしても構わない。

加算器 26 は、周波数時間変換部 24 から供給された音楽信号 $S_r(t)$ と白色雑音発生部 25 から供給された白色雑音信号 $S_w(t)$ とを時系列上で加算し、出力音楽信号 $S_o(t)$ として出力する。

なお、周波数成分 $F(f_0)$ と白色雑音信号 $S_w(t)$ に相当する周波数成分 F_w とを周波数軸上で加算した後、周波数時間変換を施して出力音楽信号 $S_o(t)$ を生成することも考えられるが、この場合、例えば日本公開特許公報平 7-221648 号や日本公開特許公報平 7-221649 号等に記載されているようなプリエコー発生等を防止する利得制御・補償手法と組み合わせた際に問題が発生する。すなわち、周波数軸上で白色雑音に相当する周波数成分 F_w を加算したとしても、その後に利得補償回路で時間軸上での利得が変化するため、白色雑音信号が生成できないという問題が発生する。このため、本実施の形態では、白色雑音は時間軸上にて生成するものとする。

以上のように、本実施の形態における音楽信号符号化装置 10 及び音楽情報復号装置 20 によれば、白色雑音成分を含む入力音楽情報を符号化する際に、符号化側において白色雑音全ての周波数成分を符号化するのではなく、白色雑音レベルのインデックス i_L や乱数インデックステーブルのインデックス i_R を符号列 C に含め、復号側において入力音楽信号の白色雑音と同等のレベルをもつ白色雑音を発生させることで、効率的な符号化を可能にすると共に、フレーム間での再生帯域の変動による雑音の発生を防止することが可能となる。

なお、本発明は、図面を参照して説明した上述の実施の形態に限定されるものではなく、添付の請求の範囲及びその主旨を逸脱することなく、様々な変更、置換又はその同等のものを行うことができることは当業者にとって明らかである。

例えば、上述の実施の形態では、ハードウェアの構成として説明したが、これに限定されるものではなく、任意の処理を、CPU (Central Processing Unit) にコンピュータプログラムを実行させることにより実現することも可能である。この場合、コンピュータプログラムは、記録媒体に記録して提供することも可能であり、また、インターネットその他の伝送媒体を介して伝送することにより提

供することも可能である。

また、上述の実施の形態では、フレーム毎の音楽信号に白色雑音が含まれる場合について説明したが、本発明は、1フレーム全体が白色雑音のみの場合にも適用可能である。この場合には、各フレームの周波数成分を分析し、

(C) 全帯域のエネルギーの分散が小さい (± 6 dB 程度)。

(D) 全帯域の周波数成分がノイズ性である。

という2つの条件を満たすフレームの平均エネルギーレベルを量子化したインデックス i_L や乱数インデックステーブルのインデックス i_R を符号列に含めるようにする。

また、白色雑音を「周波数成分」+「白色雑音レベルのインデックス i_L 及び乱数インデックステーブルのインデックス i_R 」の和として表現することも可能である。すなわち、エネルギーの大きい周波数成分からビット割当を行うことで最低限必要とされる波形再現性を保証し、エネルギーの小さい周波数成分は白色雑音レベルのインデックス i_L と乱数インデックステーブルのインデックス i_R とで置き換えることも可能である。これにより、波形再現性と符号化効率の向上とを両立させることができる。この際、ビットレートに十分な余裕があり波形再現性も必要であれば「周波数成分」に重点的にビットを配分し、ビットレートが非常に低い場合には「白色雑音レベルのインデックス i_L 及び乱数インデックステーブルのインデックス i_R 」を用いて低レート符号化を実現する、という切り替えを行うようにしても構わない。

産業上の利用可能性

上述した本発明によれば、白色雑音成分を含む音楽信号を符号化する際に、符号化側において白色雑音成分のエネルギーレベルのインデックスを符号列に含め、復号側においてその白色雑音と同等のレベルをもつ白色雑音を発生させ、復号した音楽信号と時間軸上で加算することにより、効率的な符号化を実現すると共に、ブロック間での再生帯域の変動による雑音の発生を防止することができる。

請求の範囲

1. 時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号化する音楽情報符号化装置において、

上記音楽信号中の白色雑音成分を分析する白色雑音分析手段と、

上記白色雑音分析手段によって分析した上記白色雑音成分のエネルギーレベルを表すインデックスを符号化する白色雑音符号化手段と

を備えることを特徴とする音楽情報符号化装置。

2. 請求の範囲第1項記載の音楽情報符号化装置であって、

上記白色雑音分析手段は、上記ブロック内の高域側のエネルギー分布に基づいて上記白色雑音成分を分析することを特徴とする音楽情報符号化装置。

3. 請求の範囲第1項記載の音楽情報符号化装置であって、

上記白色雑音分析手段は、上記ブロック全体のエネルギー分布に基づいて上記白色雑音成分を分析することを特徴とする音楽情報符号化装置。

4. 請求の範囲第1項記載の音楽情報符号化装置であって、

上記白色雑音符号化手段は、復号側で白色雑音成分を生成するために用いる乱数テーブルのインデックスをさらに符号化することを特徴とする音楽情報符号化装置。

5. 請求の範囲第1項記載の音楽情報符号化装置であって、

上記時間軸上の音楽信号の利得を制御する利得制御手段をさらに備えることを特徴とする音楽情報符号化装置。

6. 時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号化する音楽情報符号化方法において、

上記音楽信号中の白色雑音成分を分析する白色雑音分析工程と、

上記白色雑音分析工程にて分析した上記白色雑音成分のエネルギーレベルを表すインデックスを符号化する白色雑音符号化工程と

を有することを特徴とする音楽情報符号化方法。

7. 請求の範囲第6項記載の音楽情報符号化方法であって、

上記白色雑音符号化工程では、復号側で白色雑音成分を生成するために用いる

乱数テーブルのインデックスがさらに符号化されることを特徴とする音楽情報符号化方法。

8. 時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号化する音楽情報符号化処理をコンピュータに実行させるプログラムにおいて、

上記音楽信号中の白色雑音成分を分析する白色雑音分析工程と、

上記白色雑音分析工程にて分析した上記白色雑音成分のエネルギーレベルを表すインデックスを符号化する白色雑音符号化工程と

を有することを特徴とするプログラム。

9. 請求の範囲第8項記載のプログラムであって、

上記白色雑音符号化工程では、復号側で白色雑音成分を生成するために用いる乱数テーブルのインデックスがさらに符号化されることを特徴とするプログラム。

10. 時間軸上の音楽信号を所定の時間区間毎にブロック化し、ブロック毎に周波数変換して符号化すると共に、上記音楽信号中の白色雑音成分を分析し、該白色雑音成分のエネルギーレベルを表すインデックスを符号化して生成された符号列が記録されたことを特徴とする記録媒体。

11. 請求の範囲第10項記載の記録媒体であって、

上記符号列には、復号側で白色雑音成分を生成するために用いる乱数テーブルのインデックスがさらに符号化されて含まれていることを特徴とする記録媒体。

12. 符号化された周波数信号を復号し、逆周波数変換して時間軸上の音楽信号を生成する音楽情報復号装置において、

符号化された白色雑音成分のエネルギーレベルを表すインデックスに基づいて、時間軸上の白色雑音成分を生成する白色雑音生成手段と、

上記逆周波数変換して得られる上記時間軸上の音楽信号と上記時間軸上の白色雑音成分とを加算する加算手段と

を備えることを特徴とする音楽情報復号装置。

13. 請求の範囲第12項記載の音楽情報復号装置であって、

上記白色雑音生成手段は、符号化された乱数テーブルのインデックスに基づいて上記白色雑音成分を生成することを特徴とする音楽情報復号装置。

14. 請求の範囲第12項記載の音楽情報復号装置であって、

上記白色雑音生成手段は、符号列中の所定の値に基づいて上記白色雑音成分を生成することを特徴とする音楽情報復号装置。

15. 請求の範囲第14項記載の音楽情報復号装置であって、

上記所定の値は、正規化情報又は量子化精度情報であることを特徴とする音楽情報復号装置。

16. 請求の範囲第12項記載の音楽情報復号装置であって、

上記逆周波数変換して得られる上記時間軸上の音楽信号の利得の補償を行う利得補償手段をさらに備え、

上記加算手段は、利得補償後の上記時間軸上の音楽信号と上記時間軸上の白色雑音成分とを加算すること

を特徴とする音楽情報復号装置。

17. 符号化された周波数信号を復号し、逆周波数変換して時間軸上の音楽信号を生成する音楽情報復号方法において、

符号化された白色雑音成分のエネルギーレベルを表すインデックスに基づいて、時間軸上の白色雑音成分を生成する白色雑音生成工程と、

上記逆周波数変換して得られる上記時間軸上の音楽信号と上記時間軸上の白色雑音成分とを加算する加算工程と

を有することを特徴とする音楽情報復号方法。

18. 符号化された周波数信号を復号し、逆周波数変換して時間軸上の音楽信号を生成する音楽情報復号処理をコンピュータに実行させるプログラムにおいて、

符号化された白色雑音成分のエネルギーレベルを表すインデックスに基づいて、時間軸上の白色雑音成分を生成する白色雑音生成工程と、

上記逆周波数変換して得られる上記時間軸上の音楽信号と上記時間軸上の白色雑音成分とを加算する加算工程と

を有することを特徴とするプログラム。

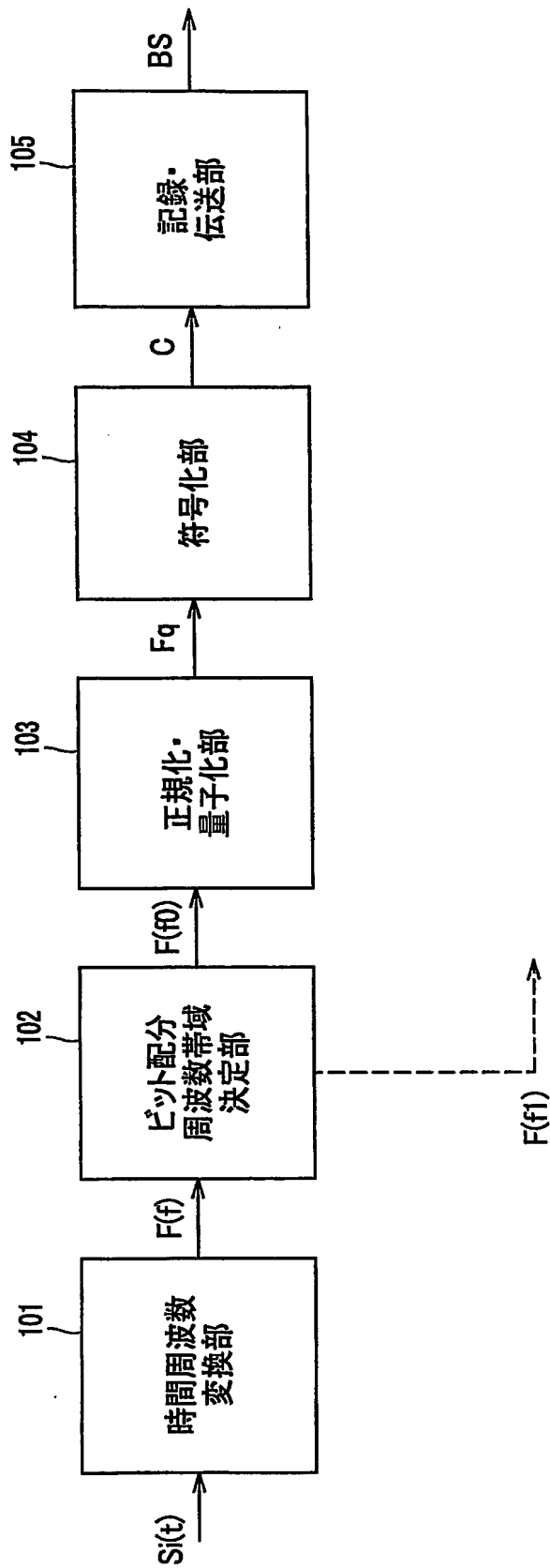


FIG.1

ヘッド H	正規化情報 SF	量子化精度情報 WL	周波数情報 SP
-------	-------------	---------------	-------------

FIG.2

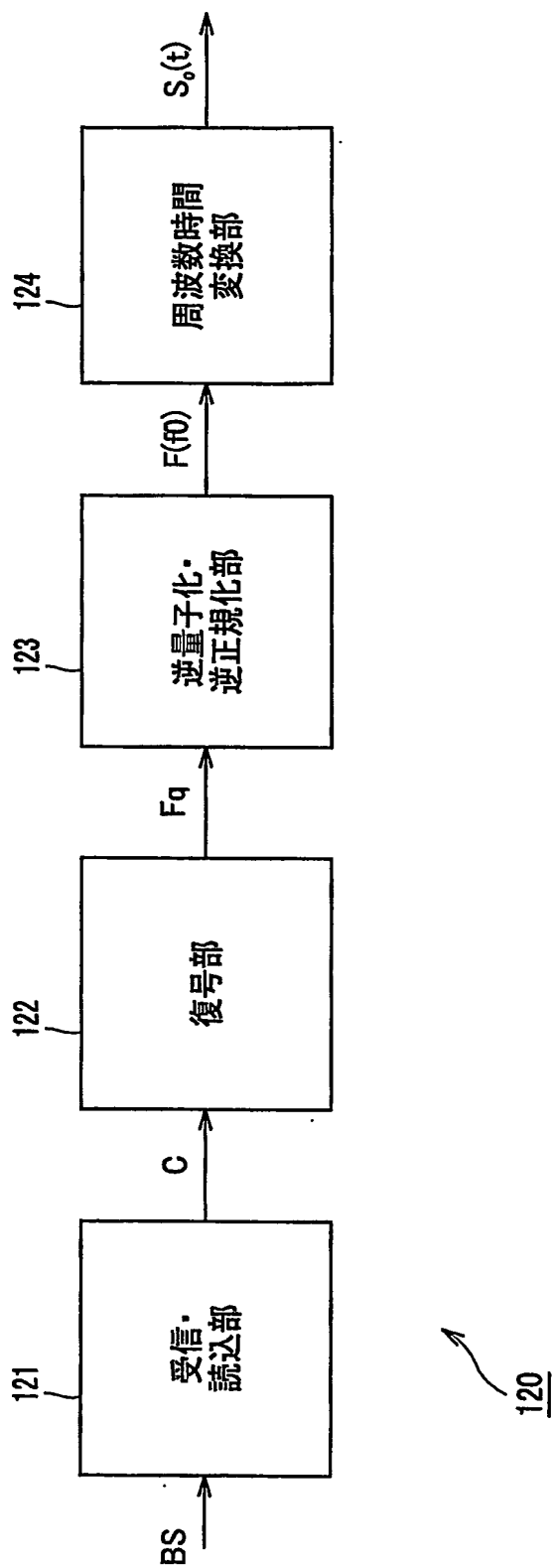


FIG.3

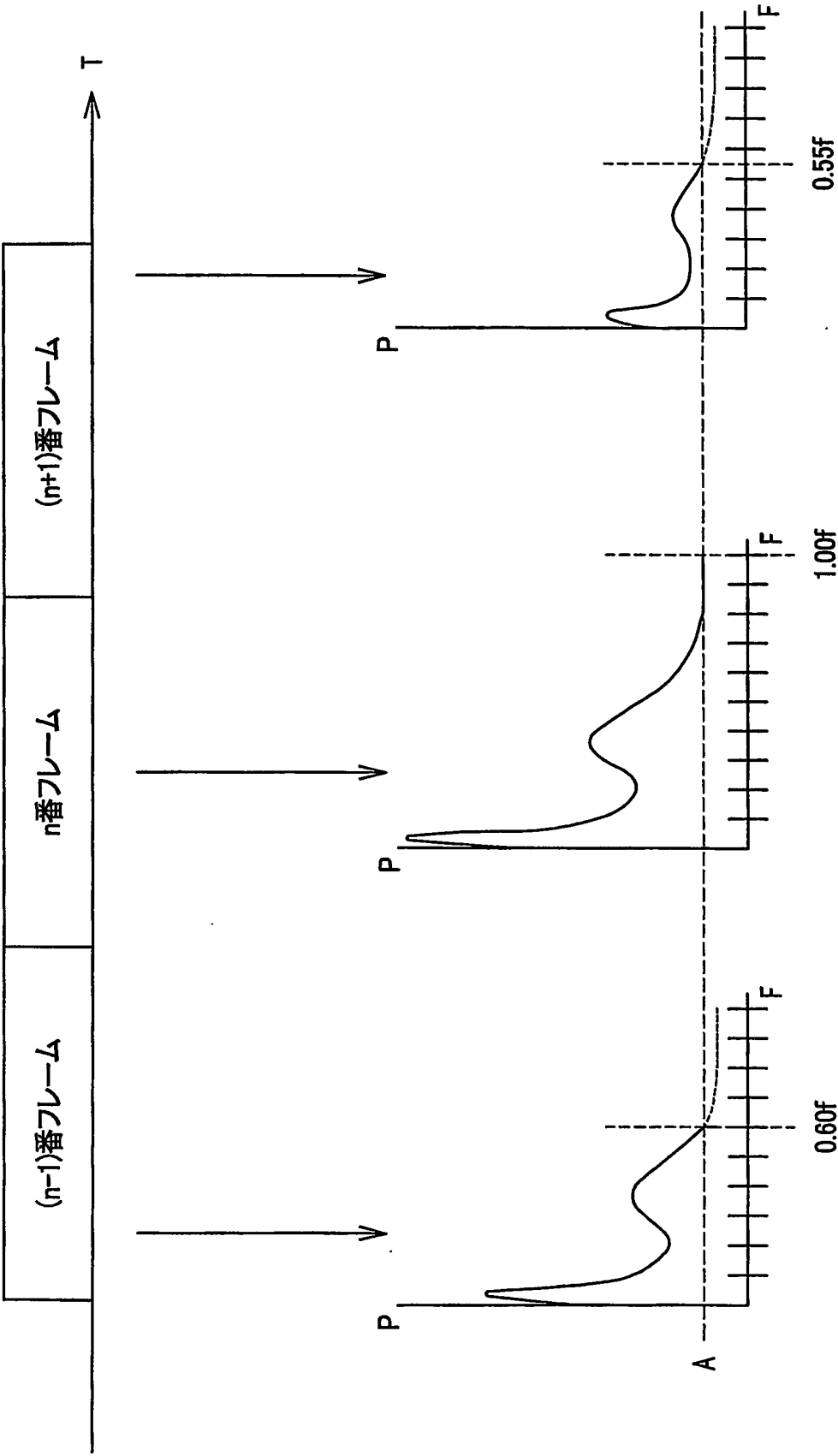


FIG.4

5/11

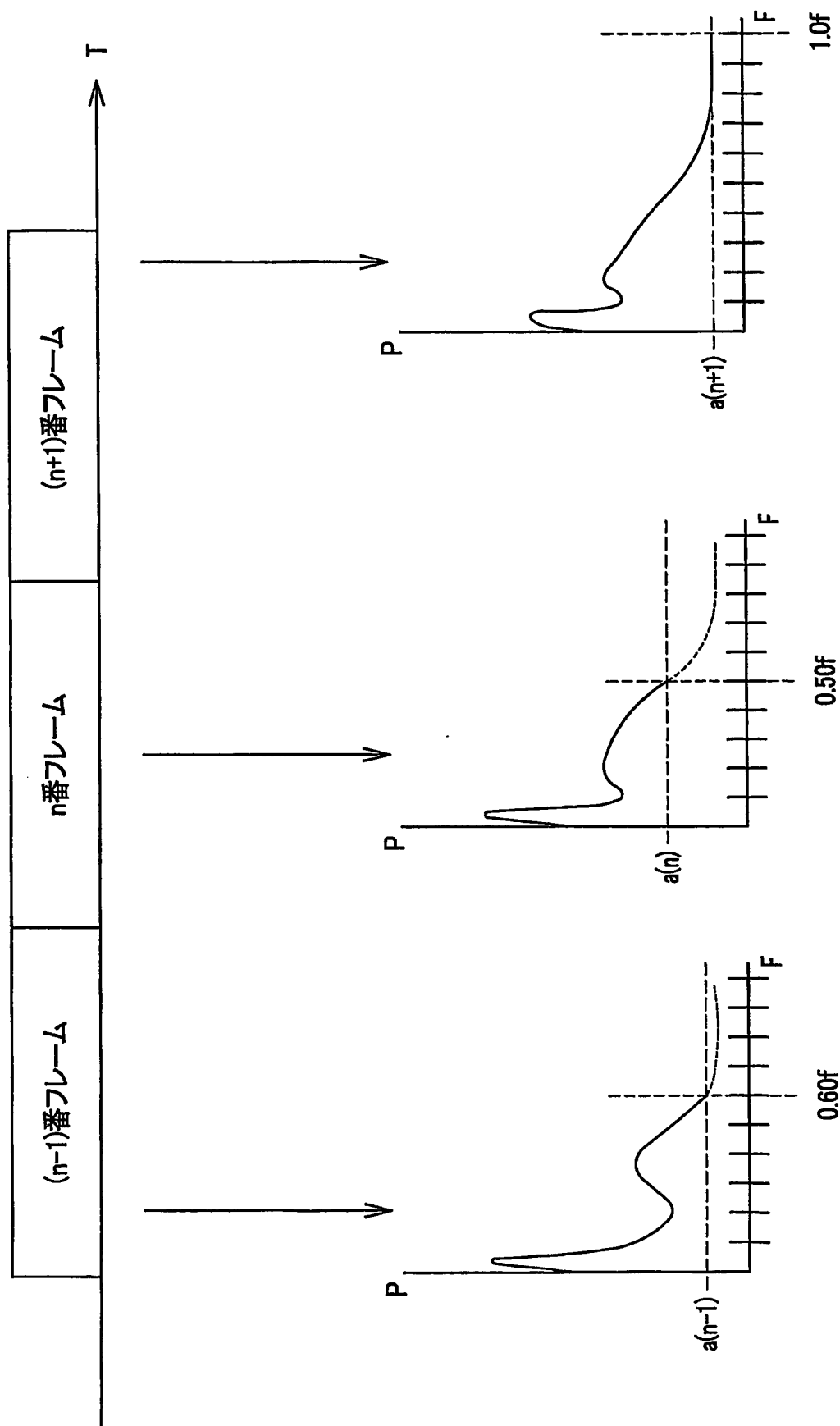


FIG.5

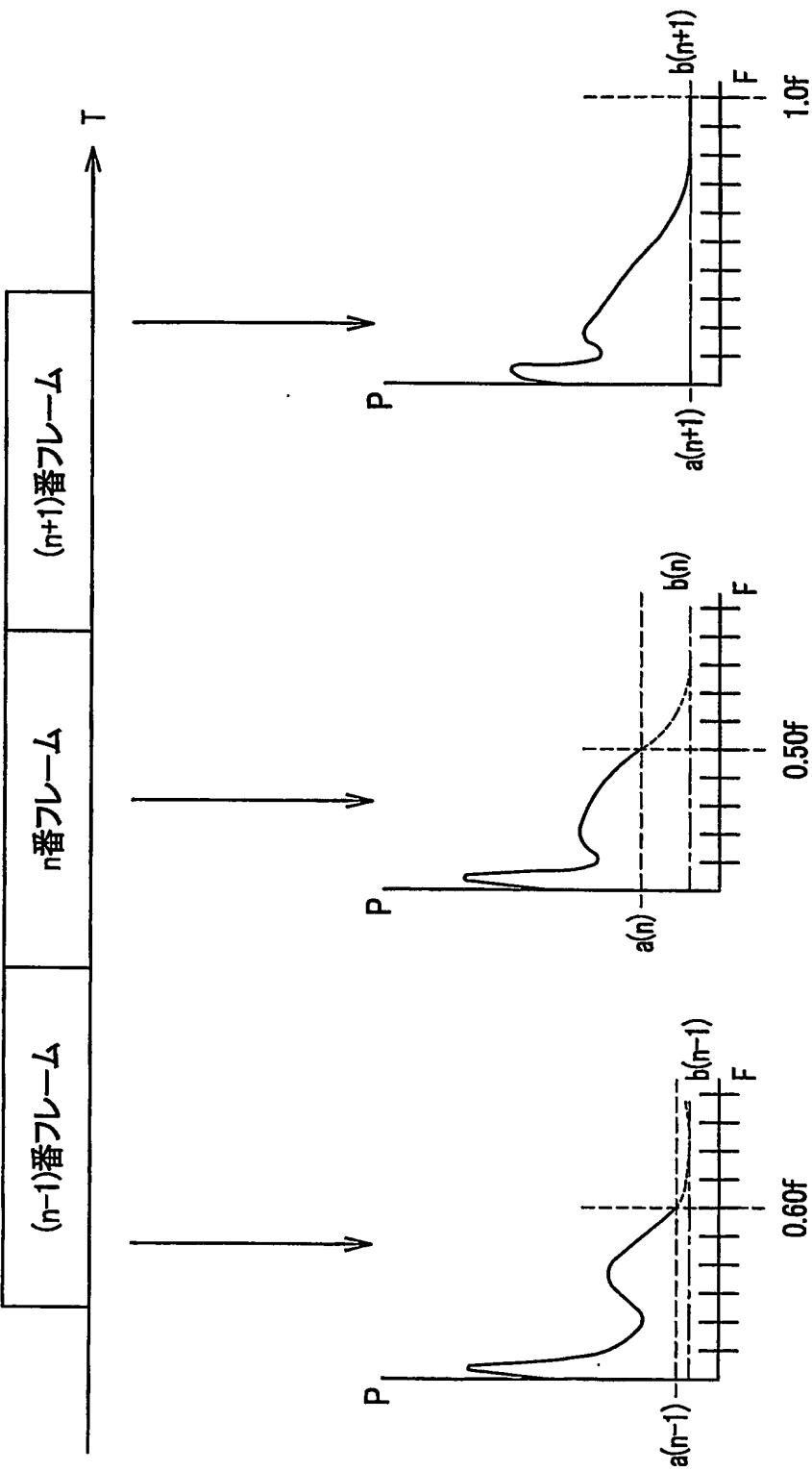


FIG.6

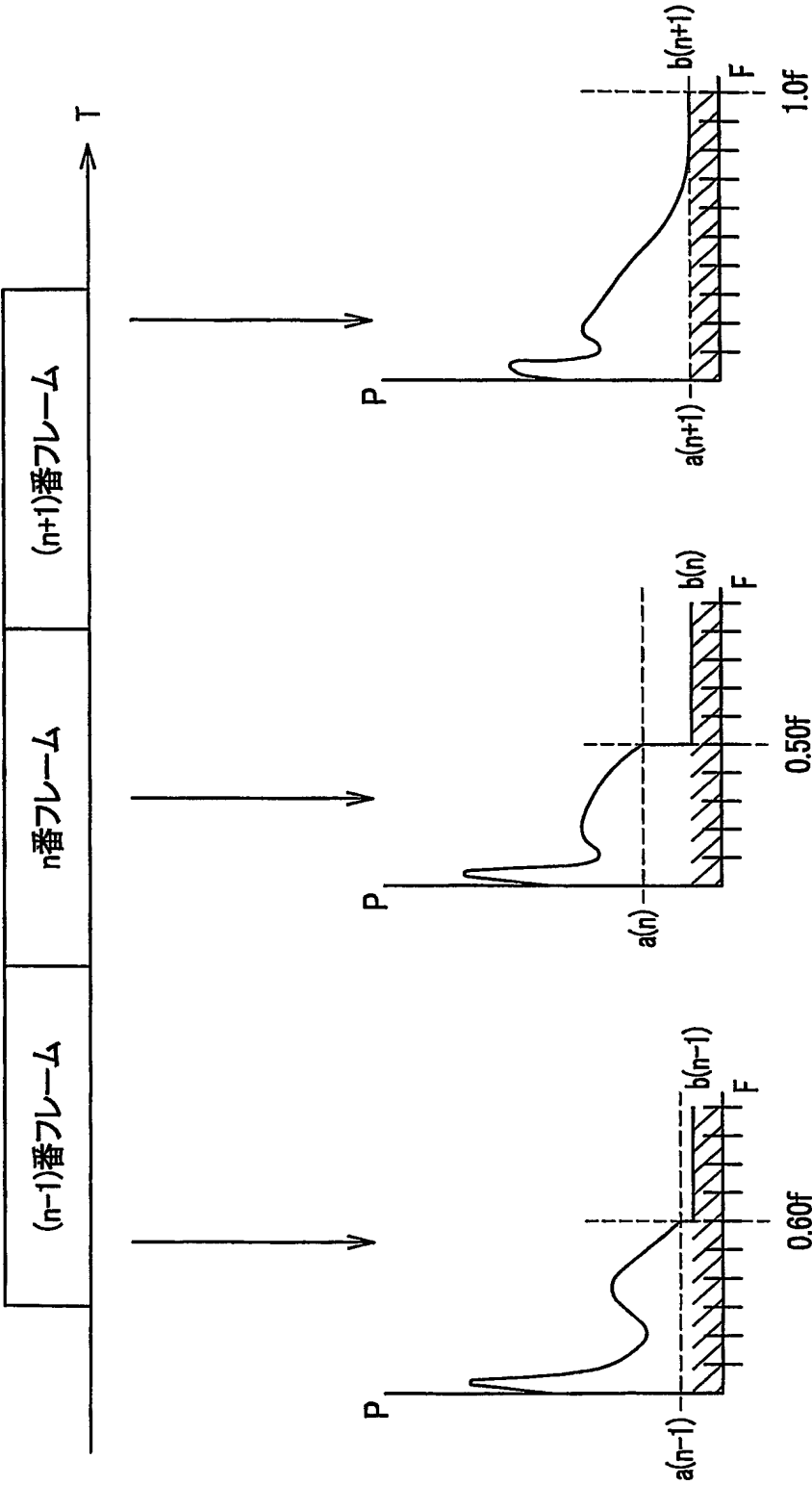


FIG.7

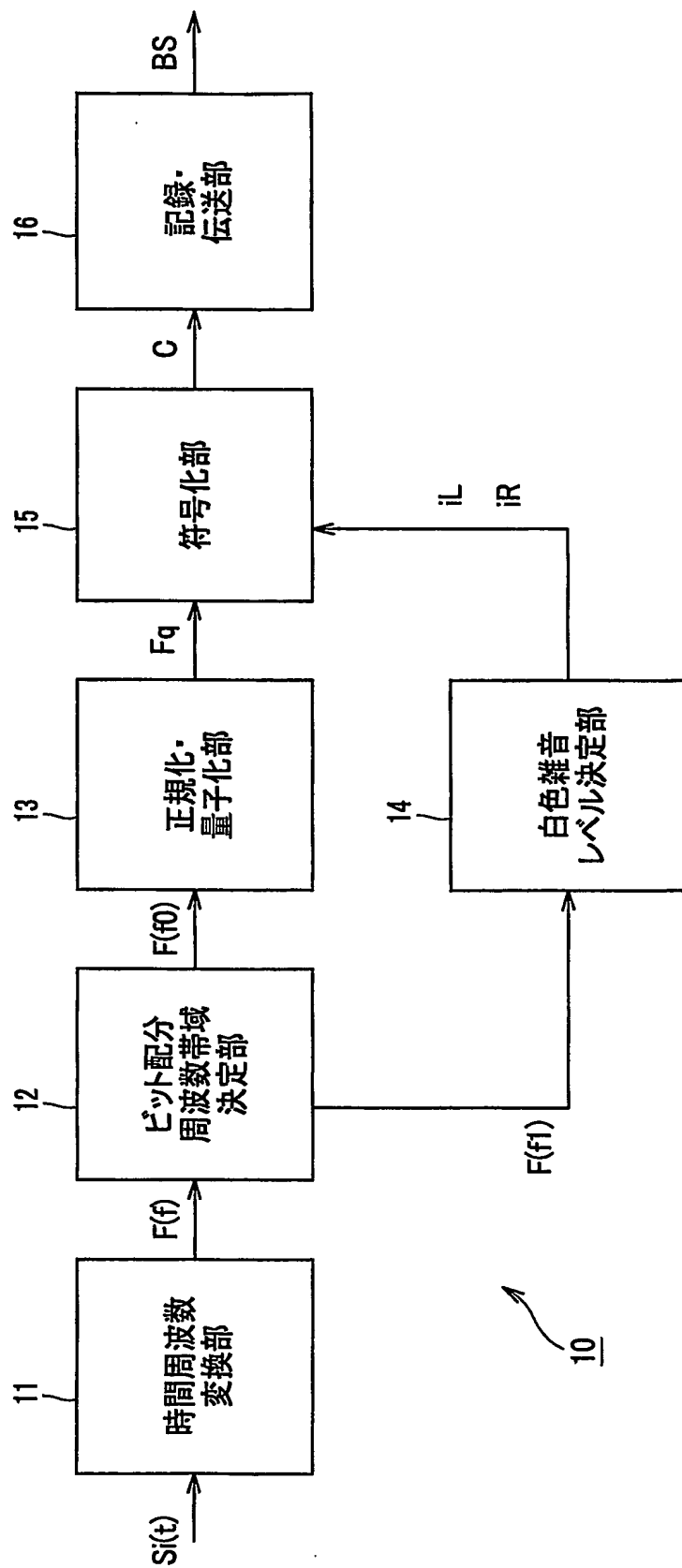


FIG.8

9/11

iL	Level
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128

FIG.9

iR	iRT
0	8
1	1005
2	912
3	63
4	567
5	281
6	736
7	459

FIG.10

ヘッダ H	正規化情報 SF	量子化精度情報 WL	周波数情報 SP	白色雑音 フラグ FL	白色雑音 情報 WN
インデックス iL			インデックス iR		

FIG.11

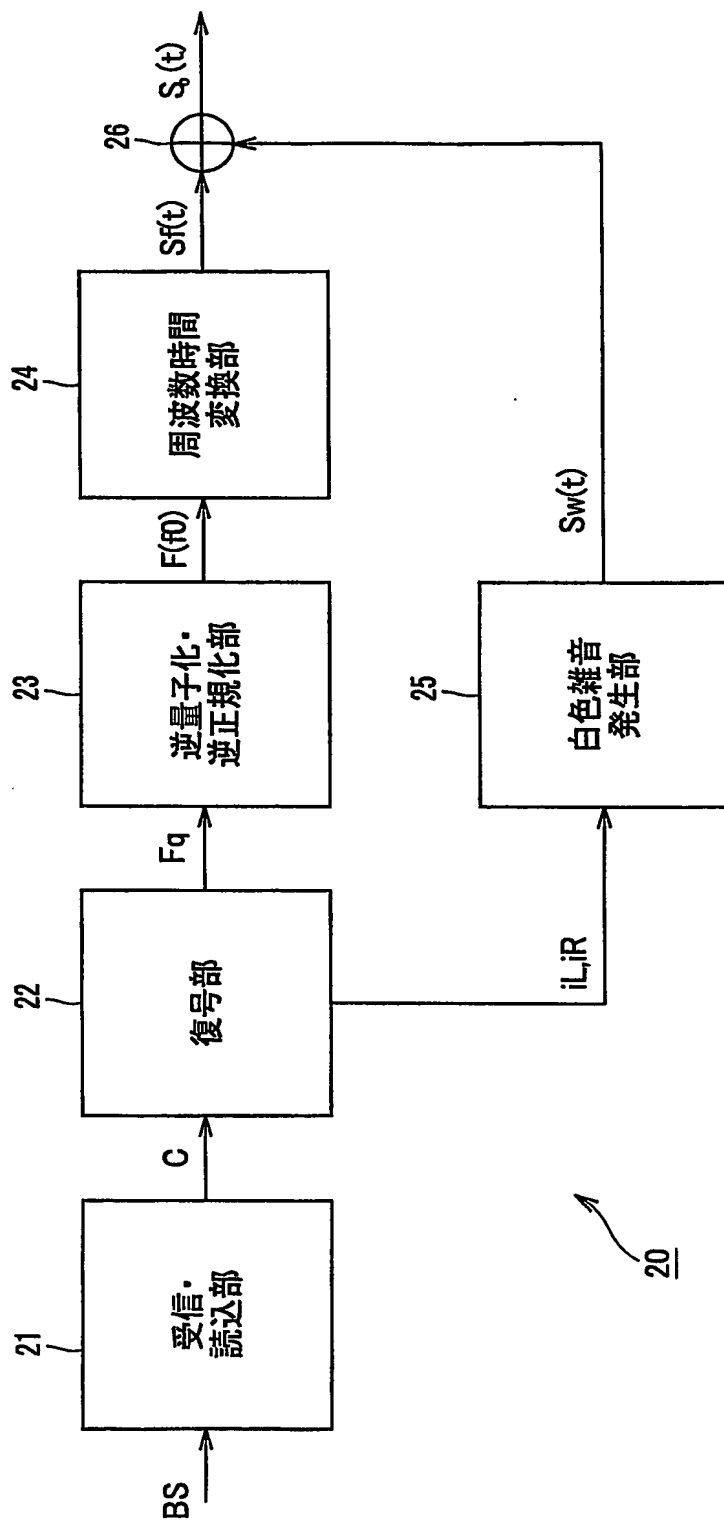


FIG.12

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/JP03/13084

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl⁷ G10L19/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 10-65546 A (Sony Corp.), 06 March, 1998 (06.03.98), & US 6097880 A1 & EP 825725 A1	1-18
Y	JP 1-28700 A (Oki Electric Industry Co., Ltd.), 31 January, 1989 (31.01.89), (Family: none)	1-18

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 31 October, 2003 (31.10.03)	Date of mailing of the international search report 18 November, 2003 (18.11.03)
--	--

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/02

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G10L19/02

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2003年
 日本国登録実用新案公報 1994-2003年
 日本国実用新案登録公報 1996-2003年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 10-65546 A (ソニー株式会社), 1998.03.06 & US 6097880 A1 & EP 825725 A1	1-18
Y	JP 1-28700 A (沖電気工業株式会社), 1989.01.31 (ファミリーなし)	1-18

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

31.10.03

国際調査報告の発送日

19.11.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号 100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)
 渡邊 聡

電話番号 03-3581-1101 内線 3540

5C 8622